ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

Vol 18, Issue 2, 2024

FAKE PROFILE IDENTIFICATION ON SOCIAL NETWORK USING MACHINE LEARNING AND NLP

[1]MR.A N L KUMAR, [2]ACHANTA VIJAYA LAKSHMI

[1](Head of the Department), MCA, Swarnandhra College

MCA, scholar, Swarnandhra College

ABSTRACT

At present social network sites are part of the life for most people. Every day several people are creating their profiles on the social network platforms and they are interacting with others independent of the user's location and time. The social network sites not only provide advantages to the users and also provide security issues to the users as well their information. To analyze who is encouraging threats in social networks we need to classify the social networks profiles of the users. From the classification, we can get the genuine profiles and fake profiles on the social networks. Traditionally, we have different classification methods for detecting fake profiles on social networks. But, we need to improve the accuracy rate of Fake profile detection in social networks. In this paper we are proposing Machine learning and Natural language Processing (NLP) techniques to improve the accuracy rate of the fake profiles detection. We can use the Support Vector Machine (SVM) and Naïve Bayes algorithm

1.INTRODUCTION

Presently, social networking has grown into a popular online pastime, drawing in hundreds of thousands of users who spend billions of minutes on such sites. There is a wide range of online social network (OSN) services available today, from those that focus on social interactions like Facebook or Myspace to those that prioritize dissemination like Twitter or Google Buzz, and even those that

bring social interaction features to existing systems like Flicker. On the other side, a major obstacle and perceived objective is expanding security measures and safeguarding OSN privacy. Every user of a social network (SN) divulges a unique amount of personal information. Our personal information is vulnerable to several forms of attack, the most serious of which might be identity theft, since it is either completely or partially exposed to the public. Theft of identity occurs when an unauthorized person exploits another person#39; s knowledge or expert witness for their own gain. Because it impacted millions of individuals all over the globe, online identity theft was a major issue in the past. Individuals who fall prey to identity theft may face a variety of consequences, such as financial loss, loss of time and money, incarceration, harm to their reputation, and loss of connections with friends and family. Many social networks now do not check the debts of regular members and have weak privacy and security regulations. The reality is that the majority of SN apps have very low levels of privacy by default, making them an ideal environment for fraud and abuse. Online social media platforms have made it easier for both experienced and inexperienced criminals to commit identity theft and impersonation assaults. Worse still, while creating an account on a social networking website, users are expected to provide accurate consent. It would be disastrous enough to lose everything if consumers#39; online activity could be easily monitored; the prospect of such bills being compromised would be devastating. Like offline profiles, online network information may be either static or dynamic. There is static knowledge, which refers to the information that a person may provide when creating a profile, and dynamic knowledge, which refers to the details that the system inside the network communicates. Differences between static and dynamic knowledge include a person#39; s hobbies and demographics as well as their runtime behaviors and network location. Current research heavily relies on both static and dynamic data. But this does ' t matter on most social networks since users only view a subset of static profiles and dynamic profiles are#39; t always evident to the person networking. For the purpose of detecting false identities and harmful information in online social networks, many methods have

been suggested by diverse researchers. Various procedures have their advantages and disadvantages. Security concerns, abuse, harassment, and trolls are just a few of the numerous issues plaguing social media. Useful for a large number of fake accounts on social media. Blank or generalized profiles are known as false profiles. These profiles include individuals who have provided fraudulent credentials. False Facebook accounts are more likely to engage in harmful and unwanted actions, which may disrupt social community consumers' experience. People make up profiles to promote and advocate for characters or groups of people, engage in social engineering, or to slander another person via online impersonation. Facebook has its own security mechanism in place to protect user credentials from various forms of spam and phishing. An analogous concept is the Facebook Immune System (FIS).More often than not, the FIS has been unable to detect Facebook user- generated bogus accounts.

2.LITERATURE SURVEY

1) Understanding User Profiles on Social Media for Fake News Detection.

AUTHOR: Kai Shu, Shuang Wang, Huan Liu – 2018

The number of people who get their news via social media is growing rapidly in recent years. Users gain from social media because of its inherent characteristics of rapid distribution, low cost, and simple access. The news is of inferior quality compared to established news leading to a significant volume of false news. As the negative impacts on people and society grow, the need of detecting false news grows. The current state of false news identification based just on content is sometimes inadequate, hence it is recommended to use user social interactions as supplementary data to enhance this area. Because of this, it is critical to comprehend the relationship between social media accounts and disinformation in great detail. This work presents the development of real-life datasets that assess the degree to which consumers trust fake news. It also comprises the selection of representative categories of users, namely experienced users who can identify falsehoods in fake news items and naïve users who are prone to believing such things. We find that these user groups may be distinguished from one another by comparing their explicit and implicit profile traits, which

can identify false news. Future research on automated false news identification may build on the results of this work.

2)Identifying Fake Profiles on LinkedIn.

AUTHOR: Featuring Shalinda Adikari and Kaushik Dutta –

There is growing importance in getting one#39; s profile visible on professional networks like LinkedIn, since more and more companies depend on these sites to create business relationships. This score is directly proportional to the desire to engage in immoral behavior on the network. 3 Building a relationship with someone whose profile is full of false material may be a huge waste of time and energy, and it can also damage the network#39; s credibility overall. True profiles might be hard to spot when they are fraudulent. Though solutions have been suggested for other social media platforms, LinkedIn accounts do not often have their associated data made public. In this study, we assess what information is essential for detecting false profiles on LinkedIn and suggest a data mining strategy for this purpose. Results acquired using bigger data sets and more extensive profile information are similar to our approach#39; s ability to detect false profiles with 87% accuracy and 89% True Negative Rate, even when working with limited profile data. Plus, our solution improves accuracy by around 14% compared to alternatives employing comparable volumes and kinds of data.

3) A Feature Based Approach to Detect Fake Profiles in Twitter.

AUTHOR:2019Jyoti Ankit Kumar Jain

The popularity of social media sites, especially Facebook and Twitter, has skyrocketed in the last decade, drawing in millions of users. Many bad actors, including spammers, have taken an interest in them because they have become a popular form of communication. Fake accounts have become more of an issue due to the increasing amount of social media users. False and fraudulent identities are heavily engaged in harmful behaviors including spamming, spreading abuse and disinformation, and artificially increasing an application#39; s users count to advocate for or influence public opinion. Therefore, it is crucial to detect these false identities in order to safeguard legitimate users from harmful intentions. We plan to combat this by using a feature-based strategy to detect these phony accounts on various social networking sites. Our effective identification of bogus accounts is based on twenty-four indicators.

4) Method for detecting spammers and fake profiles in social network;

AUTHOR: In 2019, Yuval Elo vici, Michael FIRE, and Gilad Katz

An approach to safeguarding user privacy in an online social network that uses the database of current members to choose negative examples of fraudulent accounts and positive examples of honest profiles. Next, we extract a set of traits that we know will be useful for identifying phony and real profiles by grouping the friends and followers of our instances into communities and looking at the connections inside and between them. By comparing the properties of existing fake profiles, classifiers trained using supervised learning may identify them.

5) Social Networks Fake Profiles Detection Using Machine Learning Algorithms

AUTHOR: El Yusufi twins Yasny and Zakaria - 2020

A variety of harmful actions, including advanced spersistent threats, utilize fake accounts. Finding phony accounts on social media is the main topic of this article. There are a few different ways to go about detecting false profiles on social media. One way is to look at the data associated with profiles, and another is to study individual accounts. Among all types of cybercrime, the most damaging is the establishment of false profiles on social networks. Notifying the user of the establishment of a false profile is not enough time to discover this crime. There have been several proposals in the literature for algorithms and approaches that may identify false profiles. 5 By discussing the aforementioned methods for identifying false social media accounts, this article clarifies the function of false identities in A Pts. To determine whether a profile is phony or real, we will evaluate the effectiveness of three supervised machine learning algorithms: Random Forest (RF), Decision Tree (DT-J48), and Naïve Bayes (NB).

3. EXISTING SYSTEM

Chai et al awarded on this paper is a proof-of inspiration to gain knowledge of. Even though the prototype approach has employed most effective normal systems in normal language processing and human-pc interplay, the results realized from the user trying out are significant. By comparing this simple prototype approach with a wholly deployed menu procedure, they've discovered that users, principally beginner users, strongly pick the common language dialog-based approach. They have additionally learned that

in an ecommerce environment sophistication in dialog administration is more important than the potential to manage complex typical language sentences.

In addition, to provide effortless access to knowledge on ecommerce web sites, natural language dialog-based navigation and menu-pushed navigation should be intelligently combined to meet a person's one-of-a-kind wants. Not too long ago, they accomplished development of a new iteration of the approach that includes enormous enhancements in language processing, dialog administration and information management. They believed that average language informal interfaces present powerful personalized alternatives to conventional menu pushed or search-based interfaces to web sites.

LinkedIn is greatly preferred through the folks who're in the authentic occupations. With the speedy development of social networks, persons are likely to misuse them for unethical and illegal conducts. Creation of a false profile turns into such adversary outcomes which are intricate to identify without apt research. The current solutions which were virtually developed and theorized to resolve this contention, mainly viewed the traits and the social network ties of the person's social profile. However, in relation to LinkedIn such behavioral observations are tremendously restrictive in publicly having profile data for the customers by the private insurance policies. The limited publicly available profile data of LinkedIn makes it ineligible in making use of the existing tactics in fake profile identification. For that reason, there is a distinctive study on deciding on systems for fake profile identification in LinkedIn. Shalinda Adikari and Kaushik Dutta researched and identified the minimal set of profile data that are crucial for picking out false profiles in LinkedIn and labeled the appropriate knowledge mining procedure for such a project.

Z. Halim et al. Proposed patio-temporal mining on social networks to determine the circle of customers concerned in malicious events with the support of latent semantic analysis. Then compare the results of spatial temporal coincidence with that of original organization/ties within social networks, which could be very encouraging as the organization generated by spatial-temporal co-prevalence and actual one is very nearly each other. Once they set the worth of

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

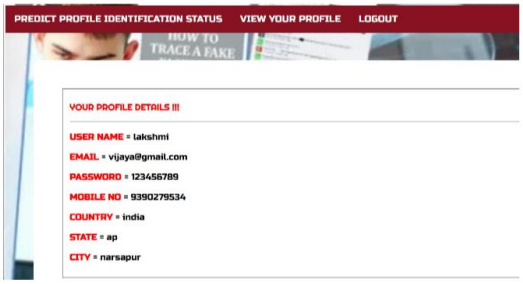
Vol 18, Issue 2, 2024

threshold to right level, we develop the number of nodes i.e. Actors so that they are able to get higher photos. Total scans indicate that Latent Semantic Indexing participates very well for picking out malicious contents, if the feature set is competently chosen. One obvious quandary of this technique is how users pick their function set and the way rich it's. If the characteristic set is very small then most of the malicious content material will not be traced. However, the bigger the person 's function set, the better the performance won.

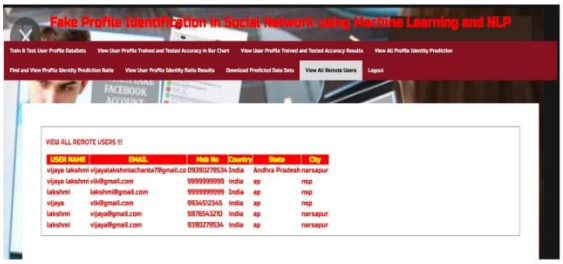
4. OUTPUT SCREENS

Home Page

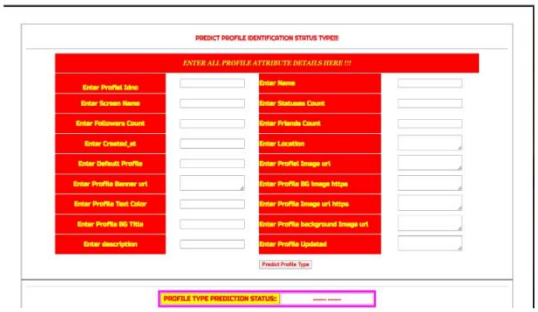
View profile page



View Remote Users



Output



5. CONCLUSION

We presented natural language processing and machine learning methods in this article. We may readily identify phony accounts on social media by using these methods. Finding the false profiles was the focus of this

research, which used the Facebook Data set. In order to examine the dataset, natural language processing (NLP) pre-processing methods are used, and machine learning algorithms like SVM and Naïve Bayes are employed for profile classification. In this study, we show that these learning techniques may increase the detection rate.

6. REFERENCE

1.Michael Fire et al. (2012). "Strangers intrusion detection-detecting spammers and fake profiles in social networks based on topology anomalies." Human Journal 1(1): 26-39. Günther, F. and S. Fritsch (2010). "Neural net: Training of neural networks." The R Journal 2(1): 30- 38

2. Dr. S. Kannan, Vaira Prakash Gurusamy, “Preprocessing Techniques for Text Mining”, 05 March 2015.

3. Shalinda Adikari and Kaushik Dutta, Identifying Fake Profiles in LinkedIn, PACIS 2014 Proceedings, AISeL.

4. Z. Halim, M. Gul, N. ul Hassan, R. Baig, S. Rehman, and F. Naz, “Malicious users’ circle detection in social network based on spatiotemporal co-occurrence,” in Computer Networks and Information Technology (ICCNIT),2011 International Conference on, July, pp. 35–390.

5. Liu Y, Gummadi K, Krishnamurthy B, Mi solves A,” Analyzing Facebook privacy settings: User expectations vs. reality”, in: Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, ACM, pp.61–70.

6. Mahmood S, Desmedt Y,” Poster: preliminary analysis of Google’s privacy. In: Proceedings of the 18th ACM conference on computer and communications security”, ACM 2011, pp.809–812.

7. Stein T, Chen E, Mangla K,” Facebook immune system. In: Proceedings of the 4th workshop on social network systems”, ACM 2011, pp 48

8. Saeed Abu-Nimeh, T. M. Chen, and O. Alzubi, “Malicious and Spam Posts in Online Social Networks,” Computer, vol.44, no.9, IEEE2011, pp.23-28

9. J. Jiang, C. Wilson, X. Wang, P. Huang, W. Sha, Yida, B. Zhao, Understanding latent interactions in online social networks, in: Proceedings of the 10th ACM SIGCOMM

Conference on Internet Measurement, ACM, 2010, pp. 369–382

10.Kazienko, P. and K. Musiał (2006). Social capital in online social networks. Knowledge-Based Intelligent Information and Engineering Systems, Springer